

CLAIMS

What is claimed is:

1. A system comprising:
a first network node and a second network node connected via a communication link;
at least one process capable of execution on said first network node;
a first monitor for said process, said first monitor capable of execution on said second network node, said monitor capable of detecting failure of said process on said first network node and causing said process to execute on said second network node.
2. The system of claim 1, wherein said first and second network nodes are central processing units.
3. The system of claim 1, wherein said first and second network nodes are computer hosts.
4. The system of claim 1, wherein said first and second network nodes are computer servers.
5. The system of claim 1, wherein said first and second network nodes are storage nodes.

6. The system of claim 1, wherein said first and second network nodes are printer nodes.

7. The system of claim 1, wherein said first and second network nodes are file systems.

8. The system of claim 1, wherein said first and second network nodes are location independent file systems.

9. The system of claim 1, wherein said communication link is a local area network.

10. The system of claim 1, wherein said communication link is a wide area network.

11. The system of claim 1, wherein said first monitor periodically checks said process executing on said first network node in order to detect a failure of said process.

12. The system of claim 11, wherein said periodic checking comprises sending a key to said process and receiving a predefined response from said process.

13. The system of claim 11, wherein said periodic checking comprises monitoring heartbeat signals sent at a periodic rate from said process.

14. The system of claim 11, wherein, when said first monitor detects the failure of said process, said first monitor initiates a process swap, said process swap comprising:

terminating said process from execution on said first network node;
initiating said process on said second network node;
initiating a second monitor on said first network node; and
terminating said first monitor from execution on said second network node.

15. The system of claim 1, wherein said process is selected from the group consisting of a service, a task and a thread.

16. A system comprising:
a first plurality of network nodes connected via a first communication link;
a second plurality of network nodes connected via a second communication link;
said first communication link and said second communication link connected through a third communication link.
a process capable of execution on one of the network nodes;

a monitor for said process capable of execution on one of the network nodes, said monitor capable of detecting failure of said process and causing said process to execute on another of the network nodes.

17. The system of claim 16, wherein said network nodes are central processing units.

18. The system of claim 16, wherein said network nodes are computer hosts.

19. The system of claim 16, wherein said network nodes are computer servers.

20. The system of claim 16, wherein said network nodes are storage nodes.

21. The system of claim 16, wherein said network nodes are printer nodes.

22. The system of claim 16, wherein said network nodes are file systems.

23. The system of claim 16, wherein said network nodes are location

independent file systems.

24. The system of claim 16, wherein said first communication link and said second communication link are local area networks.

25. The system of claim 16, wherein said third communication link is a wide area network.

26. The system of claim 16, wherein said first monitor periodically checks said process executing on said one node of said first plurality of network nodes in order to detect a failure of said process.

27. The system of claim 26, wherein said periodic checking comprises sending a key to said process and receiving a predefined response from said process.

28. The system of claim 26, wherein said periodic checking comprises monitoring heartbeat signals sent at a periodic rate from said process.

29. The system of claim 26, wherein, when said first monitor detects the failure of said process, said first monitor initiates a process swap, said process swap comprising:

terminating said process from execution;

transferring and initiating said process on another network node;
initiating a second monitor on the network node that is not the same node
as the node to which the process was transferred; and
terminating said first monitor from execution.

30. The system of claim 29, wherein, if said process initially executed
on a network node connected to said first communication link, then process
execution is initiated on a network node connected to said second communication
link.

31. The system of claim 29, wherein, if said process initially executed
on a network node connected to said second communication link, then process
execution is initiated on a network node connected to said first communication
link.

32. The system of claim 29, wherein, if said first monitor initially
executed on a network node connected to said first communication link, then
execution of said second monitor is initiated on a node connected to said second
communication link.

33. The system of claim 29, wherein, if said first monitor initially
executed on a network node connected to said second communication link, then
execution of said second monitor is initiated on a network node connected to said

first communication link.

34. The system of claim 16, wherein said process is selected from the group consists of a service, a task and a thread.

35. A method for operating a failover system, wherein failover does not require the termination of all the processes executing on a first network node, the method comprising:

executing a process on the first network node;

executing a first monitor on a second network node, said second network node connected to said first network node via a communications link;

periodically checking the operation of said process by said first monitor;

if an execution failure of said process is detected, then

terminating execution of said process on said first network node;

transferring and initiating execution of said process on said second network node;

initiating execution of a second monitor for said process on said first network node; and

terminating said first monitor.

36. The method of claim 35, wherein said first and second network nodes are a central processing units.

37. The method of claim 35, wherein said first and second network nodes are computer hosts.

38. The method of claim 35, wherein said first and second network nodes are computer servers.

39. The method of claim 35, wherein said first and second network nodes are storage nodes.

40. The method of claim 35, wherein said first and second network nodes are printer nodes.

41. The method of claim 35, wherein said first and second network nodes are file systems.

42. The method of claim 35, wherein said first and second network nodes are location independent file systems.

43. The method of claim 35, wherein said communication link is a LAN.

44. The method of claim 35, wherein said communication link is a WAN.

45. The method of claim 35, wherein said process is selected from the group consisting of a service, a task and a thread.

46. A computer system adapted to controlling failover so that the termination of all the executing processes is not required, the computer system comprising:

a first network node and a second network node;

a memory comprising software instructions adapted to enable the computer system to perform:

executing a process on said first network node;

executing a first monitor on said second network node, said second network node connected to said first network node via a communications link;

periodically checking the operation of said process by said first monitor;

if an execution failure of said process is detected, then

terminating execution of said process on said first network node;

transferring and initiating execution of said process on said second network node;

initiating execution of a second monitor for said process on said first network node; and

terminating said first monitor.

47. A computer software product for a computer system comprising a first network node and a second network node to control failover so that the termination of all the processes executing on said first network node is not required, the computer program product comprising:

software instructions for enabling the computer system to perform predetermined operations, and a computer readable medium bearing the software instructions, said predetermined operations comprising:

executing a process on said first network node;

executing a first monitor on said second network node, said second network node connected to said first network node via a communications link;

periodically checking the operation of said process by said first monitor;

if an execution failure of said process is detected, then

terminating execution of said process on said first network node;

transferring and initiating execution of said process on said second network node;

initiating execution of a second monitor for said process on said first network node; and

terminating said first monitor.

48. A method for monitoring and performing a failover of a network node connected to a communication link, the method comprising:

monitoring the operation of said network node by at least two managers;

exchanging heartbeats between said two managers;

if said first manager does not receive a heartbeat from said second manager, then said first manager executes diagnostic tests to determine how to correct the failed receipt of the heartbeat from said second manager.

49. The method of claim 48, wherein said network node is a central processing unit.

50. The method of claim 48, wherein said network node is a computer host.

51. The method of claim 48, wherein said network node is a computer server.

52. The method of claim 48, wherein said network node is a storage node.

53. The method of claim 48, wherein said network node is a printer node.

54. The method of claim 48, wherein said network node is a file system.

55. The method of claim 48, wherein said network node is a location independent file system.

56. The method of claim 48, wherein executing diagnostic tests further comprises:

attempting to access said second manager by said first manager;

attempting to access the operating system of said second manager by said first manager;

attempting to access a first network interface device of said second manager by said first manager; and

attempting to access a first switch of said second manager by said first manager.

57. The method of claim 56, wherein, if access attempt of said first network device by said first manager is unsuccessful, said first manager attempts to access said second manager through a second network interface device.

58. The method of claim 56, wherein, if access attempt of said first switch by said first manager is unsuccessful, said first manager attempts to access

said second manager through a second switch.

59. The method of claim 48, wherein determination of failure is selected from the group consisting of said second manager, a network interface device, and a switch.

60. The method of claim 57, wherein, upon determination of a failure of said first network interface device, a redundant network interface device replaces said first network interface device.

61. The method of claim 58, wherein, upon determination of a failure of said first switch, a redundant switch replaces said first switch.

62. A computer system adapted to controlling failover so that the termination of all the processes executing on a network node is not required, the computer system comprising:

a plurality of network nodes interconnected by a communication link;

a memory comprising software instructions adapted to enable the computer system to perform:

monitoring the operation of a node in the plurality of network nodes by at least two managers;

exchanging heartbeats between said two managers;

if said first manager does not receive a heartbeat from said second

manager, then said first manager executes diagnostic tests to determine how to correct the failed receipt of the heartbeat from said second manager.

62. The computer system of claim 61, wherein the software instructions adapted to executing diagnostic tests further are further adapted to:

attempt to access said second manager by said first manager;

attempt to access the operating system of said second manager by said first manager;

attempt to access a first network interface device of said second manager by said first manager; and

attempt to access a first switch of said second manager by said first manager.

63. The system of claim 62, wherein the software instructions adapted to executing diagnostic tests further are further adapted so that, if access attempt of said first network device by said first manager is unsuccessful, said first manager attempts to access said second manager through a second network interface device.

64. The system of claim 62, wherein the software instructions adapted to executing diagnostic tests further are further adapted so that, if access attempt of said first switch by said first manager is unsuccessful, said first manager attempts to access said second manager through a second switch.

65. The system of claim 63, wherein the software instructions adapted to executing diagnostic tests further are further adapted so that, upon determination of a failure of said first network interface device, a redundant network interface device replaces said first network interface device.

66. The system of claim 64, wherein the software instructions adapted to executing diagnostic tests further are further adapted so that, upon determination of a failure of said first switch, a redundant switch replaces said first switch.

67. A computer software product for monitoring and performing a failover of a network node connected to a communication link, the computer program product comprising:

software instructions for enabling the network node to perform predetermined operations, and a computer readable medium bearing the software instructions, said predetermined operations comprising:

monitoring the operation of a node in the plurality of network nodes by at least two managers;

exchanging heartbeats between said two managers;

if said first manager does not receive a heartbeat from said second manager, then said first manager executes diagnostic tests to determine how to correct the failed receipt of the heartbeat from said second manager.

68. The computer system of claim 67, wherein the predetermined operations for executing diagnostic tests further comprise:

attempting to access said second manager by said first manager;

attempting to access the operating system of said second manager by said first manager;

attempting to access a first network interface device of said second manager by said first manager; and

attempting to access a first switch of said second manager by said first manager.

69. The system of claim 68, wherein the predetermined operations for executing diagnostic tests further comprise, if access attempt of said first network device by said first manager is unsuccessful, said first manager attempts to access said second manager through a second network interface device.

70. The system of claim 68, wherein the predetermined operations for executing diagnostic tests further comprise, if access attempt of said first switch by said first manager is unsuccessful, said first manager attempts to access said second manager through a second switch.

71. The system of claim 69, wherein the predetermined operations for executing diagnostic tests further comprise, upon determination of a failure of said first network interface device, a redundant network interface device replaces said

first network interface device.

72. The system of claim 70, wherein the predetermined operations for executing diagnostic tests further comprise, upon determination of a failure of said first switch, a redundant switch replaces said first switch.